Project Proposal

Multi-source transfer learning

Fall 2019

1 Contact Information

Name	SID	Email
Jingge Wang	2019214599	wangjg19@mails.tsinghua.edu.cn
Jihong Wang	2019214581	wangjiho19@mails.tsinghua.edu.cn
Jiarong Li	2019214606	lijiaron19@mails.tsinghua.edu.cn

2 Abstract

Transfer learning has shown great power in solving learning problems when there is only a few or no labeled data in target domain in the training stage. Conventional transfer learning methods usually uses one kind of source data and use the model trained on it to solve new task on a target domain. However, in reality, we may have access to several source domain where the data all have some connection to our target task. Under such condition, we hope to use multi source data to make predictions on a new target domain. In this project, we review some well-developed multi-source domain adaptation algorithms and compare their performance. Based on the results, we also discuss the space for improvement the effect.

3 Problem Statement

(1) The core issue in transfer learning problem is: How to efficiently use the training data that are not distributed in the same way as the test data? In our case, there are more than one source domain with data distributed different as target data.

(2) The second question is: How to select appropriate source domain data from multiple source domains for knowledge transfer so as to reduce the generation of negative transfer.

(3) The final issue is to improve the existing multi-source transfer learning methods.

4 Related Work

The survey paper review some theoretical results and well-developed algorithms for the multi-source domain adaptation problem [1]. The second paper develops a method called maximal correlation weighting (MCW) to build the required target classifier from an appropriate weighting of the feature functions, thus to use information from all sources to learn a target task [2]. The third paper present an method for training the network by transferring knowledge from six publicly available texture databases and then fine-tuneing on the lung tissue data. The resulting CNNs are combined in an ensemble and their fused knowledge is compressed back to a network with the original architecture [3]. The fourth paper proposed an Incomplete Multi-source Transfer Learning framework with structured latent low-rank constraint and cross-source alignment from two directions [4]. The fifth paper propose a novel algorithm to leverage knowledge from different views and sources collaboratively, by letting different views from different sources complement each other through a co-training style framework, while revise the distribution differences in different domains [5].

5 Proposed Methodology

(1) Based on previous work, we are planning to implement several multi-source transfer learning methods to solve one real task.

(2) Evaluate the implemented methods and compare the performance of them.

(3) Propose our own modification for one of the implemented multi-source methods and hope that we can get some performance improvements over the existing ones.

6 References

[1] S. Sun, H. Shi, and Y. J. I. F. Wu, "A survey of multi-source domain adaptation," vol. 24, pp. 84-92, 2015.

[2] J. Lee, P. Sattigeri, and G. Wornell, "Learning New Tricks From Old Dogs: Multi-Source Transfer Learning From Pre-Trained Networks," in Advances in Neural Information Processing Systems, 2019, pp. 4372-4382.

[3] S. Christodoulidis, Member, IEEE, M. Anthimopoulos, M. J. I. J. o. Biomedical, and H. Informatics, "Multi-source Transfer Learning with Convolutional Neural Networks for Lung Pattern Analysis," vol. PP, no. 99, pp. 1-1, 2016.

[4] Z. Ding, H. Zhao, and Y. Fu, "Multi-source Transfer Learning," in Learning Representation for Multi-View Data Analysis: Models and ApplicationsCham: Springer International Publishing, 2019, pp. 175-202.

[5] B. Tan, E. Zhong, E. W. Xiang, and Q. Yang, "Multi-transfer: Transfer learning with multiple views and multiple sources," in Proceedings of the 2013 SIAM International Conference on Data Mining, 2013, pp. 243-251: SIAM.